MEPS HC-022:
2000 P4R3/P5R1
Population Characteristics

Agency for Healthcare Research and Quality
Center for Cost and Financing Studies

AHRQ Clearinghouse Number 01-DP06

TABLE OF CONTENTS

TABLE OF CONTENTS

A. DATA USE AGREEMENT

Individual identifiers have been removed from the micro-data contained in the files on this CD-ROM. Nevertheless, under sections 308 (d) and 903 (c) of the Public Health Service Act (42 U.S.C. 242m and 42 U.S.C. 299 a-1), data collected by the Agency for Healthcare Research and Quality (AHRQ) and/or the National Center for Health Statistics (NCHS) may not be used for any purpose other than for the purpose for which they were supplied; any effort to determine the identity of any reported cases is prohibited by law.

Therefore in accordance with the above referenced Federal Statute, it is understood that:

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and

2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director Office of Management AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity; and

3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey.

By using this data you signify your agreement to comply with the above stated statutorily based requirements with the knowledge that deliberately making a false statement in any matter within the jurisdiction of any department or agency of the Federal Government violates 18 U.S.C. 1001 and is punishable by a fine of up to $10,000 or up to 5 years in prison.

The Agency for Healthcare Research and Quality requests that users cite AHRQ and the Medical Expenditure Panel Survey as the data source in any publications or research based upon these data.

B. BACKGROUND

This documentation describes one in a series of public use files from the Medical Expenditure Panel Survey (MEPS). The survey provides a new and extensive data set on the use of health services and health care in the United States.

The Medical Expenditure Panel Survey (MEPS) is conducted to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian non-institutionalized population. MEPS also includes a nationally representative survey of nursing homes and their residents. MEPS is cosponsored by the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS).

MEPS comprises four component surveys: the Household Component (HC), the Medical Provider Component (MPC), the Insurance Component (IC), and the Nursing Home Component (NHC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. The separate NHC sample supplements the other MEPS components. Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications.

MEPS is the third in a series of national probability surveys conducted by AHRQ on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES, also known as NMES-1) was conducted in 1977, the National Medical Expenditure Survey (NMES-2) in 1987. Beginning in 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance system.

The design efficiencies incorporated into MEPS are in accordance with the Department of Health and Human Services (DHHS) Survey Integration Plan of June 1995, which focused on consolidating DHHS surveys, achieving cost efficiencies, reducing respondent burden, and enhancing analytical capacities. To accommodate these goals, new MEPS design features include linkage with the National Health Interview Survey (NHIS), from which the sampling frame for the MEPS HC is drawn, and continuous longitudinal data collection for core survey components. The MEPS HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents' medical providers, employers, and insurance providers.

1.0 Household Component

The MEPS HC, a nationally representative survey of the U.S. civilian non-institutionalized population, collects medical expenditure data at both the person and household levels. The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a 2 ½-year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for 2 calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

The sampling frame for the MEPS HC is drawn from respondents to NHIS, conducted by NCHS. NHIS provides a nationally representative sample of the U.S. civilian non-institutionalized population, with oversampling of Hispanics and blacks.

2.0 Medical Provider Component

The MEPS MPC supplements and validates information on medical care events reported in the MEPS HC by contacting medical providers and pharmacies identified by household respondents. The MPC sample includes all hospitals, hospital physicians, home health agencies, and pharmacies reported in the HC. Also included in the MPC are all office-based physicians:

- Providing care for HC respondents receiving Medicaid.
- Associated with a 75-percent sample of HC households receiving care through an HMO (health maintenance organization) or managed care plan.
- Associated with a 25-percent sample of the remaining HC households.

Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents, including:

- Diagnoses coded according to ICD-9-CM (9th Revision, International Classification of Diseases) and DSM-IV (Fourth Edition, *Diagnostic and Statistical Manual of Mental Disorders*).
- Physician procedure codes classified by CPT-4 (Common Procedure Terminology, Version 4).
- Inpatient stay codes classified by DRGs (diagnosis-related groups).

- Prescriptions coded by national drug code (NDC), medication names, strength, and quantity dispensed.

- Charges, payments, and the reasons for any difference between charges and payments.

The MPC is conducted through telephone interviews and mailed survey materials.

3.0 Insurance Component

The MEPS IC collects data on health insurance plans obtained through employers, unions, and other sources of private health insurance. Data obtained in the IC include the number and types of private insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, eligibility requirements, and employer characteristics.

Establishments participating in the MEPS IC are selected through four sampling frames:

- A list of employers or other insurance providers identified by MEPS HC respondents who report having private health insurance at the Round 1 interview.

- A Bureau of the Census list frame of private sector business establishments.

- The Census of Governments from Bureau of the Census.

- An Internal Revenue Service list of the self-employed.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and insurance providers) are linked back to data provided by the MEPS HC respondents. Data from the other three sampling frames are collected to provide annual national and State estimates of the supply of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance.

The MEPS IC is an annual panel survey. Data are collected from the selected organizations through a prescreening telephone interview, a mailed questionnaire, and a telephone followup for nonrespondents.

4.0 Nursing Home Component

The 1996 MEPS NHC was a survey of nursing homes and persons residing in or admitted to nursing homes at any time during calendar year 1996. The NHC gathered information on the demographic characteristics, residence history, health and functional status, use of services, use of prescription medications, and health care expenditures of nursing home residents. Nursing home administrators and designated staff also provided information on facility size, ownership, certification status, services provided, revenues and expenses, and other facility characteristics. Data on the income, assets, family relationships, and care-giving services for sampled nursing home residents were obtained from next-of-kin or other knowledgeable persons in the community.

The 1996 MEPS NHC sample was selected using a two-stage stratified probability design. In the first stage, facilities were selected; in the second stage, facility residents were sampled, selecting both persons in residence on January 1, 1996, and those admitted during the period January 1 through December 31.

The sample frame for facilities was derived from the National Health Provider Inventory, which is updated periodically by NCHS. The MEPS NHC data were collected in person in three rounds of data collection over a 1 ½-year period using the CAPI system. Community data were collected by telephone using computer-assisted telephone interviewing (CATI) technology. At the end of three rounds of data collection, the sample consists of approximately 815 responding facilities, 3,100 residents in the facility on January 1, and 2,200 eligible residents admitted during 1996.

5.0 Survey Management

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. NCHS provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports and microdata files. Summary reports are released as printed documents and electronic files. Microdata files are released on CD-ROM and/or as electronic files. A catalog of all MEPS products released to date is provided in Section G of this document.

Printed documents and CD-ROMs are available through the AHRQ Publications Clearinghouse. Write or call:

AHRQ Publications Clearinghouse
Attn: (publication number)
P.O. Box 8547
Silver Spring, MD 20907
800/358-9295
410/381-3150 (callers outside the United States only)
888/586-6340 (toll-free TDD service; hearing impaired only)

Be sure to specify the AHRQ number of the document or CD-ROM you are requesting. Selected electronic files are available from the Internet on the AHRQ home page: http://www.meps.ahrq.gov/.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Healthcare Research and Quality.

C. TECHNICAL AND PROGRAMMING INFORMATION

1.0 General Information

This documentation describes the fifth point in time data file to be released from the Medical Expenditure Panel Survey Household Component (MEPS HC). Released as an ASCII file (with related SAS programming statements), a SAS transport dataset, and a SAS dataset, this public use file provides information on data collected on a nationally representative sample of the civilian, non-institutionalized population of the United States during the early part of 2000. The data consist of 2000 data obtained in Round 3 of Panel 4 and Round 1 of Panel 5 of the MEPS Household Component and contains variables pertaining to survey administration, demographics, employment, health status, and health insurance.

The following documentation offers a brief overview of the types and levels of data provided, the content and structure of the files, and programming information. It contains the following sections:

- Data File Information
- Survey Sample Information
- Variable/Questionnaire Crosswalk

A codebook of all the variables included on the 2000 point in time data file is provided in a separate document (H22CB.PDF).

A copy of the MEPS survey questionnaire is included (see MEPSQUES.TXT file).

A catalog of all MEPS products released to date and a matrix indicating the major MEPS data items on public use files released to date are available on the MEPS web page (www.meps.ahrq.gov).

For more information on MEPS HC survey design, see S. Cohen, 1997; J. Cohen, 1997; and S. Cohen, 1996. For information on the MEPS MPC design, see S. Cohen, 1998.

2.0 Data File Information

This public use dataset contains variable and frequency distributions for a total of 26,018 persons (14,165 from Panel 4 Round 3 and 11,853 from Panel 5 Round 1). This count includes all household survey respondents who resided in eligible responding households. Of these persons, 25,094 were assigned a positive person level weight (13,546 from Panel 4 Round 3 and 11,548 from Panel 5 Round 1). For each variable both weighted and unweighted frequencies are provided. In conjunction with the weight variable (WGTSP13) provided on this file, data for these persons can be used to make estimates for the civilian, noninstitutionalized U.S. population as of the first half of 2000.

The records on this file can be linked to all MEPS public use data sets containing the same sample by the sample person identifier (DUPERSID). Some analysts may wish to use the 2000 data contained on this file in concert with previously released 1999 data to conduct longitudinal analysis. When using data from two MEPS public use files for the same individuals, use the most current MEPS weight and variance estimation variable to support the longitudinal comparisons.

2.1 Codebook Structure

The codebook and data file sequence lists variables in the following order:

- Unique person identifiers
- Demographic variables
- Health Status variables
- Employment variables
- Health Insurance variables
- Weight and variance estimation variables

2.2 Reserved Codes

The following reserved code values are used:

VALUE	DEFINITION
-1 INAPPLICABLE	Question was not asked due to skip pattern
-7 REFUSED	Question was asked and respondent refused to answer question
-8 DK	Question was asked and respondent did not know answer

-9 NOT ASCERTAINED	Interviewer did not record the data
-10 HOURLY WAGE >= $54.81	Variable was top-coded for confidentiality

2.3 Codebook Format

This codebook describes an ASCII data set and provides the following programming identifiers for each variable:

IDENTIFIER	DESCRIPTION
Name	Variable name (maximum of 8 characters)
Description	Variable descriptor (maximum 40 characters)
Format	Number of bytes
Type	Type of data: numeric (indicated by NUM) or character (indicated by CHAR)
Start	Beginning column position of variable in record
End	Ending column position of variable in record

2.4 Variable Naming

In general, variable names reflect the content of the variable, with an 8 character limitation. All of the variables on this file (except some demographic variables and DUID, DUPERSID, PID, and KEYNESS) end in “13” to denote they are combination Panel 5 Round 1/Panel 4 Round 3 variables. For edited variables the “13” is followed by an “X” and are so noted in the variable label. Variables contained in this delivery were derived either from the questionnaire itself or from the CAPI. The source of each variable is identified in the section of the documentation entitled “E. Variable-Source Crosswalk”. Sources for each variable are indicated in one of four ways: (1) variables derived from CAPI or assigned in sampling are so indicated; (2) variables derived from complex algorithms associated with re-enumeration are labeled “RE Section”; (3) variables that come from one or more specific questions have those numbers listed in the “Source” column; (4) variables constructed from multiple questions using complex algorithms are labeled “Constructed” in the “Source” column.

2.5 File Contents

2.5.1 Survey Administration Variables

The survey administration variables contain information related to conducting the interview, household and family composition, and person-level and RU-level status codes. Data for the survey administration variables were derived from the sampling process, the CAPI programs, or were computed based on information provided by the respondent in the reenumeration section of the questionnaire. Most Survey Administration variables on this file are asked during every round of the MEPS interview. Variables in this delivery describe data for Panel 4 Round 3 and Panel 5 Round 1 in 2000.

The variable PANEL13 indicates the panel from which the data are derived. A value of 4 indicates Panel 4 Round 3 data and a value of 5 indicates Panel 5 Round 1 data.

Note that Round 3 of Panel 4 covers both the end of 1999 and the beginning of 2000. (When possible, the variables were constructed to represent data from the 2000 portion of Round 3.)

Dwelling Units, Reporting Units, and Families

The definitions of Dwelling Units (DUs) in the MEPS Household Survey is generally consistent with the definition employed for the National Health Interview Survey. The dwelling unit ID (DUID) is a five-digit random ID number assigned after the case was sampled for MEPS. The person number (PID) uniquely identifies all persons within the dwelling unit. The variable DUPERSID is a combination of the variables DUID and PID, thus uniquely identifies each sampled person in MEPS.

A Reporting Unit (RU) is a person or group of persons in the sampled dwelling unit who are related by blood, marriage, adoption, foster care or other family association. Each RU is to be interviewed as a single entity for MEPS. Thus, the RU serves chiefly as a family-based "survey operations" unit rather than an analytic unit. Members of each RU within the DU are identified by the variable RULETR13. Regardless of the legal status of their association, two persons living together as a "family" unit were treated as a single reporting unit if they chose to be so identified. Examples of different types of reporting units are:

1. A married daughter and her husband living with her parents in the same dwelling unit constitute a single reporting unit.

2. A husband and wife and their unmarried daughter, age 18, who is living away from home while at college constitute two reporting units.

3. Three unrelated persons living in the same dwelling unit would each constitute a distinct reporting unit, three reporting units in all.

Unmarried college students less than 24 years of age who usually live in the sampled household, but were living away from home and going to school at the time of the MEPS interview, were treated as a Reporting Unit separate from that of their parents for the purpose of data collection. The variable RUSIZE13 indicates

the number of persons in each RU, treating each student as a single RU separate from their parents. Thus, students are not included in the RUSIZE13 count of their parents' RU. However, for many analytic objectives, the student reporting units would be combined with their parents' reporting unit, treating the combined entity as a single family. Family identifier and size variables are described below and include students with their parents' reporting unit.

The variable FAMID13 identifies a family (i.e., persons living together related to one another by blood, marriage, adoption, foster care, or self-identified as a single unit plus related students who are living away at post-secondary school) for each round. These family identifier variables use a letter and a DU identifier to indicate a person's family affiliation. In order to identify a person's family affiliation, users must create a unique family identification variable by concatenating the DU identifier (DUID) and the FAMID13 variable, as described in Section 3.2.2.3 Instructions to Create Family Estimates.

The variable FAMSIZ13 indicates the number of persons associated with a single family unit after students are linked to their associated parent RUs for analytical purposes. Family-level analyses should use the FAMSIZ13 variables. In a few cases, students were deleted from the file because attempts to contact them were unsuccessful, and no data were collected for them. However, these persons are accounted for in the FAMSIZ13 variable.

The family size (FAMSIZ13) and the reporting unit (RU) size (RUSIZE13) counts may not be consistent with the count of records on the file. There are some reporting units where the RU size variable (RUSIZE13) is not equal to the number of people in that RU actually included on the file. This occurs because people who did not respond for their entire period of eligibility were not included on the file. In addition, for a number of these reporting units, the reference person is not included on the file for this same reason.

The variable RURSLT13 indicates the RU response status for Round 3 for the Panel 4 sample and Round 1 for the Panel 5 sample. The values include the following:

60 Complete with RU member
61 Complete with proxy--all RU members deceased on or after 1/1/2000
62 Complete with proxy--all RU members institutionalized or deceased on or after 1/1/2000
63 Complete with proxy, other

There are several other variables that characterize the reporting unit. The variable RUCLAS13 indicates the RU classification. RUs are classified for fielding purposes as 1 "Standard", 2 "New RU", or 3 "Student RU". Standard RUs are the original RUs from NHIS. A new RU is one which has been created when members of the household leave the standard RU and are followed according to the rules of the survey. A student RU is one in which an unmarried college student under 24 years of age is considered a usual member of the household but was living away from home while going to school and was treated as a Reporting Unit (RU) separate from that of their parents for the purpose of data collection.

Reference Period Dates

The reference period is the period of time for which data were collected in each round for each person. The reference period dates were determined during the interview for each person by the CAPI program.

The round-specific beginning reference period dates are included for each person. These variables include BEGRFM13, BEGRFD13, and BEGRFY13. The reference period for Panel 5 Round 1 for most persons identified at NHIS began on January 1, 2000 and ended on the date of the Round 1 interview. Persons who joined the RU after 1/1/2000 have their beginning reference date for the round as the day they joined the RU.

For Panel 4 Round 3 the reference period for most persons began on the date of the previous round's interview and ended on the date of the current round's interview. Persons who joined after the previous round's interview had their beginning reference date for the round set as the day they joined the RU. Persons who were present only for the 1999 portion of Round 3 are not included in this delivery.

The dates of the interview and the ending reference period dates are included for each person. These variables include RUENDM13, RUENDD13, RUENDY13, ENDRFM13, ENDRFD13, and ENDRFY13. In general, the date of the interview is the reference period end date for most persons. Note that the end date of the reference period is prior to the date of the interview if the person was deceased during the round, left the country, was institutionalized prior to that round's interview, or joined the military during the round and was not living with someone else who was eligible. If a person left the RU and that person was key and in-scope, the person was followed in the new RU to which he or she moved and his or her reference period dates pertain to the new RU.

Reference Person Identifiers

The variable RNDREF13 identifies the reference person for the RU. In general, the reference person is defined as the household member 16 years of age or older who owns or rents the home. If more than one person meets this description, the household respondent identifies one from among them. If the respondent was unable to identify a person fitting this definition, the questionnaire asked for the head of household and this person was then considered the reference person for that RU. This information was collected in the reenumeration section of the CAPI questionnaire.

Respondent Identifiers

The respondent is the person who answered the interview questions for the reporting unit (RU). The round-specific variable RDRESP13 identifies the respondent. Only one respondent is identified for each RU. In instances where the interview was completed in more than one session, only the first respondent is indicated.

There are two types of respondents. The respondent can be either an RU member or a non-RU member proxy. The variable PROXY13 identifies the type of respondent.

Person Status

A number of variables describe the various components reflecting each person's status for each round of data collection. These variables provide information about a person's in-scope status, keyness status, eligibility

status, and disposition status. These variables include: INSCOP13, KEYNESS, and PSTAT13. These variables are set based on sampling information and responses provided in the reenumeration section of the CAPI questionnaire.

Through the reenumeration section of the CAPI questionnaire, each member of a reporting unit was classified as "key" or "non-key", "in-scope" or "out-of-scope", and "eligible" or "ineligible" for MEPS data collection. To be included in the set of persons used in the derivation of MEPS person level estimates, a person had to be a member of the civilian non-institutionalized population for at least one day during 2000. Because a person's eligibility for the survey might have changed since the NHIS interview, a reenumeration of household membership was conducted at the start of each round's interview. Only persons who were "in-scope" sometime during 2000, "key", and responded for the full period in which they were in-scope were assigned person level weights and thus are to be used in the derivation of person level national estimates from the MEPS.

In-Scope

A person is considered as in-scope during a round if he or she is a member of the U.S. civilian, non-institutionalized population at some time during that round. The variable INSCOP13 indicates a person's in-scope status, specifically indicating whether a person was ever in-scope during the 2000 portion of the round.

Keyness

The term "keyness" is related to an individual's chance of being included in MEPS for purposes of making estimates about the U. S. civilian non-institutionalized population. A person is key if that person is linked for sampling purposes to the set of NHIS sampled households designated for inclusion in MEPS. Specifically, a key person either was a member of an NHIS household at the time of the NHIS interview, or was a family member who began living with a member of such a household after being out-of-scope prior to joining that member. (Examples of the latter situation include newborns and persons returning from military service, an institution, or living outside the United States.)

A non-key person is one whose chance of selection for the NHIS (and MEPS) was associated with a household eligible but not sampled for the NHIS, and who later became a member of a MEPS reporting unit. MEPS data (e.g., utilization and income) were collected for the period of time a non-key person was part of the sampled unit to provide information for family level analyses. However, non-key persons who leave a sample household unaccompanied by a key, in-scope member were not followed for subsequent interviews. Non-key individuals do not receive person level sample weights and thus do not contribute to person level national estimates. They may receive family level weights if they are a member of a responding family.

The variable KEYNESS indicates a person's keyness status. This variable is not round-specific. Instead, it is set at the time the person enters MEPS, and the person's keyness status never changes. Once a person is determined to be key, that person will always be key.

It should be pointed out that a person may be key even though not part of the civilian, non-institutionalized portion of the U.S. population. For example, a person in the military may have been living with his or her

civilian spouse and children in a household sampled for the NHIS. The person in the military would be considered a key person for MEPS. However, such a person would not be eligible to receive a person-level sample weight if he or she was never in-scope during 2000. He or she may receive a family weight if a member of a responding family.

Eligibility

The issue of a person's eligibility for MEPS is a data collection issue. Data are to be collected only for persons considered eligible for MEPS.

All key, in-scope persons of a sampled RU are eligible for data collection. The only non-key persons eligible for data collection are those who happen to be living in an RU with at least one key, in-scope person. Their eligibility continues only for the time that they are living with at least one such person. The only out-of-scope persons eligible for data collection are those persons serving full-time on active duty in the military who were living with key in-scope persons, and again only for the time they are living with such a person.

A person may be classified as eligible for an entire round or for some part of a round. For persons who are eligible for only part of a round, data are collected for that person only for the period of time for which that person was classified as eligible.

Person Disposition Status

The variable PSTAT13 indicates a person's response and eligibility status. The PSTAT13 variable indicates the reasons for either continuing data collection for a person or terminating data collection for each person in the MEPS. Using this variable, one could identify persons who moved during the reference period, died, were born, were institutionalized or were in the military.

The following codes specify the value labels for the PSTAT13 variables. Note that some values for PSTAT13 are round-specific, as indicated in the labels.

11 Person in original RU, not full-time active military duty

12 Person in original RU, full-time active military duty, out-of-scope for whole reference period

13 Person is a full-time student living away from home, but associated with sampled RU

14 Person is full-time active military duty during round and is in-scope for part of the reference period and is in the RU at the end of the reference period

22 Person leaves a health care institution and rejoins the community - round 3 only

23 Person leaves a health care institution, goes into community and then dies - round 3 only

31 Person from original RU, dies during reference period

32 Person entered health care institution during reference period

33 Person entered non-health care institution during reference period

34 Person moved from original RU, outside US (not as student)

35 Person moved from original RU, to a military facility while on full-time active military duty

41 Person moved from the original RU, to new RU within US (new RUs include RUs originally classified as "Student RU" but which converted to "New RU")

42 Person joins RU and is not full-time military during round or joins RU and is in the military the entire round

44 Person leaves an RU and joins an existing RU and is not both in the military and coded as in-scope during the round

51 Person is newborn in reference period

Geographic Variables

The variable REGION13 indicates the Census region for the RU. MSA13 indicates whether or not the RU is found in a metropolitan statistical area. These variables indicate the geographic location of the reporting unit. The region variable is coded according to the Census regions, and the MSA13 variable reflects the June 30, 1993 definition of metropolitan statistical areas.

2.5.2 Demographic Variables

These variables provide information about the demographic characteristics of each person. As noted below, some variables have edited and imputed values. Values of most demographic variables on this file are obtained during each round of the MEPS interview. These variables describe data for Panel 4, Round 3 and Panel 5, Round 1, as well as a number of characteristics that are not round specific.

Sex

The variable SEX contains data on the sex of each RU member (SEX), as determined during the NHIS interview; it was verified and, if necessary, corrected during each MEPS interview. The data for new RU members (persons who were not members of the RU at the time of the NHIS interview) were also obtained during each MEPS Round. When sex of the RU member was not available from the NHIS interview and was not ascertained during one of the subsequent MEPS interviews, it was assigned in the following way. The person's first name was used to assign sex, if obvious. If the person's first name provided no indication of

gender, then family relationships were reviewed. If neither of these approaches made it possible to determine the individual's sex, sex was randomly assigned.

Age

Date of birth and age for each RU member were asked or verified during each MEPS interview (DOBMM, DOBYY, AGE13X). If date of birth was available, age was calculated based on the difference between date of birth and date of interview. Inconsistencies between the calculated age and the age reported during the CAPI interview were reviewed and resolved. For purposes of confidentiality, the variable AGE13X was top coded at 90 years, and DOBYY bottom coded at 1910. When date of birth was not provided but age was (from either the MEPS or the NHIS data), the month and year of birth were assigned randomly from among the possible valid options. For any cases still not accounted for, age was imputed using (1) the mean age difference between MEPS participants with certain family relationships (where available) or (2) the mean age value for MEPS participants. For example, a mother's age is imputed as her child's age plus the mean age difference between MEPS mothers and their children, or a wife's age is imputed as the husband's age plus the mean age difference between MEPS wives and husbands.

Race, Race/Ethnicity, Hispanic Ethnicity, and Hispanic Ethnicity Group

Race (RACEX) and Hispanic ethnicity (HISPANX) were asked for each RU member during the MEPS interview. If this information was not obtained in Round 1, the questions were asked in subsequent Rounds. When race and/or ethnicity was not reported in the interview, values for these variables were obtained based on the following priority order. When available, they were obtained from the originally collected NHIS data. If not ascertained, the race, and/or ethnicity were assigned based on relationship to other members of the DU using a priority ordering that gave precedence to blood relatives in the immediate family. The variable RACETHNX indicating both race and ethnicity (e.g., with categories such as "Hispanic" and "black but not Hispanic") reflects the imputations done for RACEX and HISPANX. The specific Hispanic ethnicity group is given in the unedited variable HISPCAT.

Student Status and Educational Attainment

The variable FTSTD13X indicates whether the person was a full-time student at the interview date. This variable has valid values for all persons between the ages of 17 - 23 inclusive.

The variables indicating completed years of education when first entered MEPS (EDUCYR13) and highest degree when first entered MEPS (HIDEG13) were obtained from questions RE 103-105. Starting with Round 3 of Panel 4, and Round 1 of Panel 5 questions RE103-105 were asked only when persons first entered MEPS, which was Round 1 for most people.

For the completed years of education variable (EDUCYR13), children who are 5 years of age or older when they first entered MEPS and who never attended school were coded as 0; children under the age of 5 years were coded as -1 "Inapplicable" regardless of whether or not they attended school.

The highest degree (HIDEG13) was obtained from two questions: high school diploma (RE 104) and highest degree (RE 105). Persons under 16 years of age when they first entered MEPS were coded as 8 "Under 16 - Inapplicable". In cases where the response to the highest degree question was "no degree" and highest grade was 13 through 17, the variable was coded as 3 "high school diploma". If highest grade completed for those with a "no degree" response was "refused" or "don't know", the variable was coded as 1 "no degree". The user should note that the EDUCYR13 and HIDEG13 variables are unedited variables and minimal data cleaning was performed on these variables. Therefore, discrepancies in data may remain for these two sets of variables. Decisions as to how to handle these discrepancies are left to the analyst.

Marital Status and Spouse ID

Current marital status was collected and/or updated during each Round of the MEPS interview. This information was obtained in RE13 and RE97 and is reported as MARRY13X. Persons under the age of 16 were coded as 6 "under 16 - inapplicable". In instances where there were discrepancies between the marital status of two individuals within a family, other person-level variables were reviewed to determine the edited marital status for each individual. For example, in Panel 4 Round 3, when one spouse was reported as married and the other spouse reported as widowed, the data were reviewed to determine if one partner should be coded as 8 "widowed in Round".

Four edits were performed to ensure minimal consistency across rounds for the Panel 4 Round 3 data. First, a person could not be coded as "Never Married" after previously being coded as any other marital status (e.g. "Widowed"). Second, a person could not be coded as "Under 16 - Inapplicable" after being previously coded as any other marital status. Third, a person could not be coded as "Married in Round" after being coded as "Married" in the Round immediately preceding. Fourth, a person could not be coded as an "in Round" code (e.g., "widowed in Round") in two subsequent Rounds.

When marital status was missing in the preceding round and provided in the current round, then the person was coded to the "in round" marital state. For example, if marital status was not available from the National Health Interview Survey, and the person's marital status was reported as married in round 1 of MEPS, then the person would be coded as "7 married in round" for round 1 of MEPS.

The person identifier for each individual's spouse is reported in SPOUID13. The variable is set to the PID (within each family) of the person identified as the spouse during the round. If no spouse was identified in the household, the variable was coded as 995 "no spouse in house". Those with unknown marital status are coded as 996 "marital status unknown". Persons under the age of 16 are coded as 997 "Less than 16 years old".

The SPOUIN13 variable indicates whether a person's spouse was present in the RU during the Round. If the person had no spouse in the household, the value was coded as 2. For persons under the age of 16 the value was coded as 3. The SPOUID13 and SPOUIN13 variables were obtained from RE76 and RE77, where the respondent was asked to identify how each pair of persons in the household was related. Analysts should note that this information was collected in a set of questions separate from the questions that asked about marital status. While editing was performed to ensure that SPOUID13 and SPOUIN13 are consistent within each Round, there was no consistency check between these variables and marital status in a given Round. Apparent discrepancies between marital status and spouse information may be due to any of the following

causes: 1) Ambiguity as to when during a Round a change in marital status occurred. This is a result of relationship information being asked for all persons living in the household at any time during the Round, while marital status is asked as of the interview date (e.g., If one spouse died during the reference period, the surviving spouse's marital status would be "widowed in Round", but SPOUIN13 and SPOUID13 for the same round would indicate that a spouse was present); 2) Valid discrepancies in the case of persons who are married but not living with their spouse, or separating but still living together; or 3) Discrepancies which cannot be explained for either of the previous reasons.

Military Service and Service Era

Information on active duty military status was collected during each Round of the MEPS interview. Persons currently on full-time active duty status are identified in the variable ACTDTY13. Those under 16 years of age were coded as 3 "under 16- inapplicable" and those over the age of 59 were coded as 4 "over 59- inapplicable".

Relationship to the Reference Person within Reporting Units

For each reporting unit (RU), the person who owns or rents the dwelling unit is usually defined as the reference person. For student RUs, the student is defined as the reference person. (For additional information on reference persons, see the documentation on Reference Person Identifiers in the survey administration section.) The variable RFREL13X indicates the relationship of each individual to the reference person of the reporting unit (RU) in a given round. For the reference person, this variable has the value "self"; for all other persons in the RU, relationship to the reference person is indicated by codes representing "husband/spouse," "wife/spouse," "son," "daughter," "female partner," "male partner," etc. A code of 91, meaning "other related," was used to indicate rarely observed relationship descriptions such as "mother of partner. If the relationship of an individual to the reference person was not ascertained during the Round-specific interview, relationships between other RU members were used, where possible, to assign a relationship to the reference person. If MEPS data were not sufficient to identify the relationship of an individual to the reference person, relationship variables from the NHIS data were used to assign a relationship. In the event that a meaningful value could not be determined or data were missing, the relationship variable was assigned a missing value code.

2.5.3 Health Status Variables

Health Status variables involved the construction of person-level variables based on information collected in the Condition Enumeration and Health Status sections of the questionnaire. The majority of Health Status questions were initially asked at the family level to ascertain if anyone in the household had a particular problem or limitation. These were followed up with questions to determine which household member had each problem or limitation. Logical edits were performed in constructing the person-level variables to assure that family-level and person-level values were consistent. Particular attention was given to cases where missing values were reported at the family level to ensure that appropriate information was carried to the person level. Inapplicable cases occurred when a question was never asked because of a skip patterns in the

survey (e.g., individuals who were 13 years of age or older were not asked some follow-up verification questions). Inapplicable cases are coded as -1. In addition, for all variables, deceased persons were coded as inapplicable and received a code of -1.

Perceived Health Status and Mental Health Status

Perceived health status (RTHLTH13) and mental health status (MNHLTH13) were collected in the Condition Enumeration section. These questions (CE01 and CE02) asked the respondent to rate each person in the family according to the following categories: excellent, very good, good, fair, and poor. No editing was done to these variables. The corresponding dichotomous variables RTPROX13 and MNPROX13 each indicate whether the ratings of physical and mental health, respectively, were provided by oneself or by someone else.

IADL and ADL Help/Supervision

The Instrumental Activities of Daily Living (IADL) Help or Supervision variable (IADLHP13) was constructed from a series of three questions. The initial question (HE01) determined if anyone in the family received help or supervision with IADLs such as using the telephone, paying bills, taking medications, preparing light meals, doing laundry, or going shopping. If the response was "yes", a follow-up question (HE02) was asked to determine which household member received this help or supervision. For persons under age 13, a final verification question (HE03) was asked to confirm that the IADL help or supervision was the result of an impairment or physical or mental health problem. If the response to the final verification question was "no", IADLHP13 was coded as "no" for persons under the age of 13.

If no one in the family was identified as receiving help or supervision with IADLs, all members of the family were coded as receiving no IADL help or supervision. In cases where the response to the family-level question was "don't know", "refused", or otherwise missing, all persons were coded according to the family-level response. In cases where the response to the family-level question (HE01) was "yes" but no specific individuals were identified in the follow-up question as having IADL difficulties, all persons were coded as "don't know" (-8).

The Activities of Daily Living (ADL) Help or Supervision variable (ADLHLP13) was constructed in the same manner as IADLHP13, but using questions HE04-HE06. Coding conventions for missing data were the same as for IADLHP13.

Functional Limitations

A series of questions pertained to functional limitations, defined as difficulty in performing certain specific physical actions. WLKLIM13 was the filter question. It was derived from a question (HE09) that was asked at the family level: Does anyone in the family have difficulties walking, climbing stairs, grasping objects, reaching overhead, lifting, bending or stooping, or standing for long periods of time? If the answer was no then all family members were coded as "no" (2) on WLKLIM13. If the answer was "yes", then the specific persons who had any of these difficulties were identified and coded as "yes" (1) on WLKLIM13, and

remaining family members were coded as "no". If the response to the family-level question was "don't know" (-8), "refused" (-7), "missing" (-9), or "inapplicable" (-1), then the corresponding missing value code was applied to each family members value for WLKLIM13. If the answer to HE09 was "yes", but no specific individual was named as experiencing such difficulties, then each family member was assigned -8 for WLKLIM13. Deceased respondents were assigned a -1 code ("inapplicable") for WLKLIM13.

If any family member was coded "yes" to WLKLIM13, a subsequent series of questions was administered. The series of questions for which WLKLIM13 served as a filter was as follows:

LFTDIF13 - difficulty lifting 10 pounds
STPDIF13 - difficulty walking up 10 steps
WLKDIF13 - difficulty walking 3 blocks
MILDIF13 - difficulty walking a mile
STNDIF13 - difficulty standing 20 minutes
BENDIF13 - difficulty bending or stooping
RCHDIF13 - difficulty reaching over head
FNGRDF13 - difficulty using fingers to grasp

The series of questions was asked separately for each person who was coded "yes" to WLKLIM13. The series of questions was not asked for other individual family members for whom WLKLIM13 was "no". In addition, this series was not asked about family members who were less than 13 years of age, regardless of their status on WLKLIM13. Finally, these questions were not asked about deceased family members. In such cases (i.e., WLKLIM13 = 2, or age < 13, or PSTATS13 = 31), each question in the series was coded as "inapplicable" (-1). Finally, if responses to WLKLIM13 were "refused" (-7), "don't know" (-8), "not ascertained" (-9), or otherwise inapplicable (-1), then each question in this series was coded as "inapplicable" (-1).

Analysts should note that, for WLKLIM13, there was no minimum age criterion that was used to determine a skip pattern, whereas, for the subsequent series of questions, persons less than 13 years old were skipped and coded as inapplicable. Therefore, it is possible for someone aged 12 or less to have a code of 1 yes on WLKLIM13, and also to have codes of inapplicable on the subsequent series of questions.

Use of Assistive Technology and Social/Recreational Limitations

The variables indicating use of assistive technology (AIDHLP13, from question HE07) and social/recreational limitations (SOCLIM13, from question HE22) were collected initially at the family level. If there was a "yes" response to the family-level question, a second question identified the specific individual(s) to whom the "yes" response pertained. Each individual identified as having the difficulty was coded "yes" on the appropriate variable; all remaining family members were coded "no". If the family-level response was "don't know", "refused", or otherwise missing, all persons were coded with the family-level response. In cases where the family-level response was "yes" but no specific individual was identified as having difficulty, all family members were coded as "don't know".

Work, Housework, and School Limitations

The variable indicating any limitation in work, housework, or school (ACTLIM13) was constructed using questions HE19-HE20. Specifically, information was collected initially at the family level. If there was a "yes" response to the family-level question (HE19), a second question (HE20) identified the specific individual(s) to whom the "yes" response pertained. Each individual identified as having a limitation was coded "yes" on ACTLIM13; all remaining family members were coded "no". If the family-level response was "don't know", "refused", or otherwise missing, all persons were coded with the family-level response. In cases where the family-level response was "yes" but no specific individual was identified as having difficulty, all family members were coded as "don't know" (-8). Persons less than five years old were coded as "inapplicable" (-1) on ACTLIM13.

If ACTLIM13 was "yes" and the person was 5 years of age or older, a follow-up question (HE20A) was asked to identify the specific limitation or limitations for each person. These included working at a job (WRKLIM13), doing housework (HSELIM13), or going to school (SCHLIM13). Respondents could answer "yes" to each activity; one person could thus report limitation in multiple activities. WRKLIM13, HSELIM13, and SCHLIM13 have values of "yes" or "no" only if ACTLIM13 was "yes"; each variable was coded as "inapplicable" (-1) if ACTLIM13 was "no", "refused", or otherwise missing. When ACTLIM13 was "don't know", these variables were all coded as "don't know". If a person was under 5 years old or was deceased, WRKLIM13, HSELIM13, and SCHLIM13 were each coded as "inapplicable" (-1).

A second question (UNABLE13) asked if the person was completely unable to work at a job, do housework, or go to school. This question was asked only of the same set of respondents who provided data on WRKLIM13, HSELIM13, and SCHLIM13. Therefore, those respondents who were coded "no" on ACTLIM13, or were under 5 years of age, or were deceased, were coded as "inapplicable" (-1) on UNABLE13. UNABLE13 was asked once for whichever set of WRKLIM13, HSELIM13, and SCHLIM13 the respondent had limitations; if a respondent was limited in more than one of these three activities, UNABLE13 did <u>not</u> specify if the respondent was completely unable to perform all of them, or only some of them.

Cognitive Limitations

The variable (COGLIM13) was collected at the family level as a three-part question (HE24-01 to HE24-03) indicating if any of the adults in the family (1) experience confusion or memory loss, (2) have problems making decisions, or (3) require supervision for their own safety. If a "yes" response was obtained to any item, the persons affected were identified in HE25 and COGLIM13 was coded as "yes". Remaining family members not identified were coded as "no" for COGLIM13.

If responses to HE24-01 though HE24-03 were all "no", or if two of three were "no" and the remaining was "don't know", "refused", or otherwise missing, all family members were coded as "no". If responses to the three questions were combinations of "don't know", "refused", and missing, all persons were coded as "don't know". If the response to any of the three questions was "yes" but no individual was identified in HE25, all persons were coded as "don't know".

COGLIM13 reflects whether any of the three component questions is "yes". Respondents with one, two, or three specific cognitive limitations cannot be distinguished. In addition, because the question asked specifically about "adult" family members, all persons less than 18 years of age are coded as "inapplicable" (-1) on this question.

2.5.4 Employment Variables

Employment questions were asked of all persons 16 years and older at the time of the interview. Employment variables consist of person-level indicators such as employment status and job-related variables such as hourly wage. All job-specific variables refer to a person's current main job. The current main job, defined by the respondent, indicates the main source of employment.

Employment variables included on the Panel 4 Round 3/Panel 5 Round 1 2000 release are: EMPST13, HRWAG13X, HRWGRD13, HRWAY13, HOUR13, HELD13X, OFFER13X, NUMEMP13 and SELFCM13. Most employment variables pertain to status as of the date of the interview.

With the exception of health insurance held or offered from a current main job, no attempt has been made to logically edit any employment variables. When missing, values were imputed for certain persons' hourly wage; however, there was no editing performed on any values reported by the respondent. Hourly wages greater than or equal to $54.81 were top-coded to –10. The number of employees variable was top-coded at 500.

Employment Status (EMPST13)

Employment status was asked for all persons aged 16 or older. Responses to the employment status question were: "currently employed" if the person had a job at the interview date, "has a job to return to" if the person did not work during the reference period but had a job to return to as of the interview date, "employed during the reference period" if the person had no job at the interview date but did work during 2000, and "not employed with no job to return to" if the person did not have a job at the interview date, did not work during the reference period, and did not have a job to return to. These responses are mutually exclusive. A current main job was defined for persons reporting that they were currently employed and who identified a current main job, and for persons who reported and identified a job to return to. Therefore, job-specific information such as hourly wage exists for persons not presently working at the interview date but who have a job to return to.

Hourly wage (HRWAG13X, HRWGRD13, and HRWAY13)

Hourly wage was asked of all persons who reported a current main job that was not self-employment (SELFCM13). For reasons of confidentiality, the hourly wage variable (HRWAG13X) was top-coded. A value of –10 indicates that the hourly wage was greater than or equal to $54.81 The hourly wage on this file (HRWAG13X) should be considered along with its accompanying variables HRWGRD13 and HRWAY13. HRWGRD13 is a flag that indicates the round in which the reported hourly wage was collected. This flag is

always set to "1" for people who are a part of Panel 5 because the reported hourly wage is always from Round 1 as only Round 1 information is reported on this file. People who are a part of Panel 4 can have a current main job from a previous round and HRWGRD13 indicates the round in which the wage information was collected. For Round 3 current main jobs that continue as the current main job from Round 1, HRWGRD13 is "1". For Round 3 current main jobs that continue as the current main job from Round 2 (but not Round 1), HRWGRD13 is "2". For Round 3 current main jobs that are identified as current main for the first time in Round 3, HRWGRD13 is "3".

For persons who did not indicate a wage amount but who did indicate a range into which the hourly wage falls, the reported hourly wage (HRWAG13X) is the median within that range. The medians were calculated using actual wages reported from the same round by persons of the same gender reporting hourly wages within each age range category. In some cases, particularly in the low wage range, gender was not used in the calculation of the median wage in order to provide a large enough base.

HRWAY13 indicates how the corresponding HRWAG13X was constructed. Hourly wage was derived, as applicable, from a large number of source variables. In the simplest case, hourly wage was reported directly by the respondent. For other persons, construction of the hourly wage was based upon their salary, the time period on which the salary was based, and the number of hours worked per time period. If the number of hours worked per time period was not available, a value of 40 hours per week was assumed, as identified in the HRWAY13 variable.

Health Insurance (HELD13X and OFFER13X)

There are two employment-related health insurance measures included in this release: health insurance held from a current main job (HELD13X) and health insurance offered from a current main job (OFFER13X). The held and offer variables were logically edited using health insurance information not available for public release.

HELD13X is "yes" if the person has a current main job where the person is not self-employed with firm size = 1, reports insurance from the employer or union at that job, and this coverage provides hospital/physician benefits or Medigap benefits. HELD13X is also "yes" if the person's current main job is with the armed forces. HELD13X is "no" if the person does not hold a current main job with the armed forces, is not self-employed at the current main job, and either reported that health insurance is not provided through that job or reported insurance but then disavowed it. To disavow insurance is to initially report it but then to deny that it is provided later in the interview or to confirm it but to indicate that it does not include hospital/physician benefits or Medigap benefits.

OFFER13X is "yes" if HELD13X is "yes" or if person has a current main job where person is not self-employed with firm size = 1 and insurance was offered through the employer or union at that job. OFFER13X is "no" if HELD13X is "no" and if the person has a current main job where person is not self-employed with firm size = 1 and insurance was not offered by the employer or union at that job.

As indicated above, information collected in the health insurance section of the interview was considered in the construction of HELD13X and OFFER13X. For example, several persons indicated in the employment section of the interview that they held health insurance through a current main job and then denied this coverage later in the health insurance section. Such people were coded as "no" for HELD13X. Due to

questionnaire skip patterns, the value for HELD13X was considered in constructing the OFFER13X variable. For example, if a person responded that health insurance was held from a current main job, they were skipped past the question relating to whether health insurance was offered at that job. If the person later disavowed this insurance in the health insurance section of the questionnaire, we would not be able to ascertain whether they were offered a policy. These individuals are coded as -9 for OFFER13X.

Finally, persons under age 16 as well as persons aged 16 and older who did not hold a current main job or who were self-employed with no employees were coded as inapplicable for the health insurance-related employment variables.

Hours (HOUR13)

HOUR13 is the number of hours worked per week.

Number of Employees (NUMEMP13)

Due to confidentiality concerns, the variable indicating the number of employees at the establishment (NUMEMP13) has been top coded at 500 or more employees. NUMEMP13 indicates the number of employees at the location of the person's current main job. For persons who reported a categorical size, we report a median estimated size from within the reported range.

2.5.5 Health Insurance Variables

Constructed and edited variables are provided for general categories of health insurance coverage collected during the MEPS Panel 5/Round 1 and Panel 4/Round 3 interviews. These variables include CHNOW13X (CHAMPUS/CHAMPVA/TRICARE coverage), MCARE13 (unedited Medicare coverage), MCARE13X (edited Medicare coverage), OTPUB13X (other public coverage including Medicaid and other government hospital/physician coverage), PRIV13 (private health insurance coverage), and INSRD13X (any health insurance coverage). With the exception of PRIV13, the insurance variables for the Panel 5/Round 1 observations have been edited. For the Panel 5/Round 1 sample, minimal editing was performed on the Other Public Coverage and Medicare variables to assign persons to coverage from these sources. For CHAMPUS/CHAMPVA/TRICARE coverage, persons who were over age 65 had their reported CHAMPUS/CHAMPVA/TRICARE coverage overturned. As mentioned above, private insurance coverage was unedited and unimputed for Panel 5/Round 1. For Panel 4/Round 3, most of the insurance variables have been logically edited to address issues that arose during Rounds 2 and 3 when reviewing insurance reported in earlier rounds. One edit corrects for possible respondent confusion with respect to a question about covered benefits asked of respondents who reported a change in their private health insurance plan name. Additional edits were performed to address issues of missing data on the time period of coverage. Note that the Medicare and CHAMPUS/CHAMPVA/TRICARE variables indicate coverage at the time of the Panel 5/Round 1 or Panel 4/Round 3 interview dates. The private coverage and other public insurance variables indicate coverage at any time during Panel 5/Round 1 or Panel 4/Round 3.

Medicare

Medicare (MCARE13) coverage was edited (MCARE13X) for persons age 65 or over. Within this age group, individuals were assigned Medicare coverage if:

> They answered yes to a follow-up question on whether or not they received Social Security benefits; or
>
> They were covered by Medicaid, other public hospital/physician coverage or Medigap coverage; or
>
> Their spouse was age 65 or older and covered by Medicare; or
>
> They reported CHAMPUS/CHAMPVA/TRICARE coverage.

Other Public Coverage

The other public coverage variable (OTPUB13X) refers to coverage both by Medicaid and to other public hospital/physician coverage. The MEPS questionnaire asks respondents about Medicaid coverage and then asks a follow-up question on other public hospital/physician coverage in an attempt to identify Medicaid recipients who may not have recognized their coverage as Medicaid. These questions were asked only if a respondent did not report having Medicaid coverage. The variable OTPUB13X is set to yes if a respondent indicated coverage from Medicaid or other public hospital/physician coverage.

Private Insurance

This file includes a variable indicating whether a household respondent was covered by private insurance at any time during the early part of 2000 (PRIV13). Private insurance could have been obtained from an employer, union or have been purchased directly either as part of a group or as non-group coverage. Private health insurance coverage was also reported where the respondent could not identify the source of the coverage or the coverage was obtained through a policyholder outside the household. An individual was considered to have private health insurance coverage if, at a minimum, that coverage provided benefits for hospital and physician services (including Medigap coverage). Sources of insurance with missing information regarding the type of coverage were assumed to contain hospital/physician coverage. Persons without private hospital/physician insurance were not counted as privately insured.

Any Insurance in Round 3 / Round 1

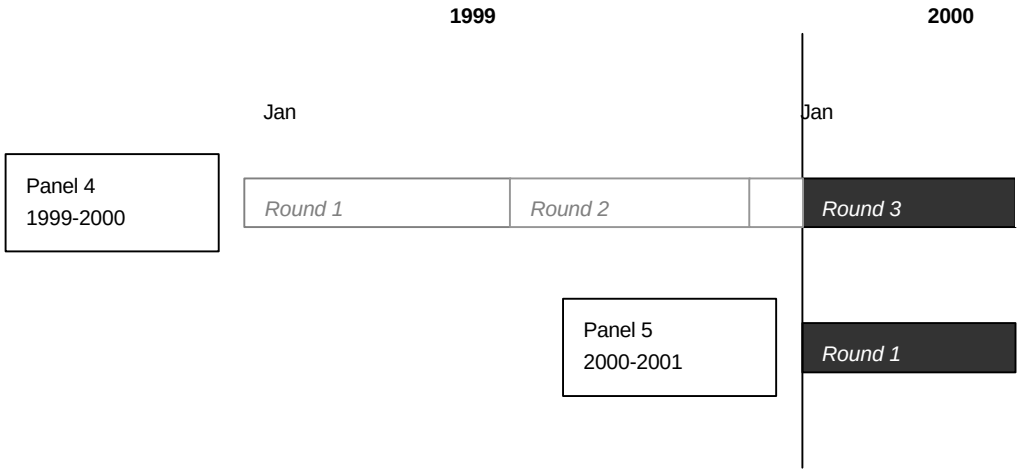
The file also includes a summary measure that indicates whether or not a sample person has any insurance during the early part of 2000 (INSRD13X). Persons identified as insured are those reporting coverage under CHAMPUS/CHAMPVA/TRICARE, Medicare, Medicaid or other public hospital/physician or private

hospital/physician insurance (including Medigap plans). A person is considered uninsured if not covered by one of these insurance sources.

3.0 Survey Sample Information

3.1 Sample Design and Response Rates

The MEPS is designed to produce estimates at the national and regional level over time for the civilian non-institutionalized population of the United States and some subpopulations of interest. Data are collected for each MEPS panel to cover a two-year period, with the two MEPS panels spanning 1999-2000 and 2000-2001, respectively. As described previously, this file consists of the subset of data from the fourth and fifth MEPS panels covering approximately the first half of calendar year 2000. More specifically, data from the 2000 portion of the third round of data collection for the MEPS Panel 4 sample are pooled with data from the first round of data collection for the MEPS Panel 5 sample (illustrated below).



3.1.1 Panel 4

For MEPS Panel 4 6,875 households were fielded in 1999, a nationally representative subsample of the households responding to the 1998 National Health Interview Survey (NHIS). Similar to the earlier MEPS panels, the Panel 4 sample reflects the oversampling of Hispanics and Black households resulting from the NHIS sample design. Hispanic households were oversampled at a rate of roughly 2 to 1, while the oversampling rate for Black households was roughly 1.5 to 1.

The overall MEPS Panel 4 response rate at the end of round 3 (when data were collected for the first part of 2000) was 65.2 percent. This overall rate reflects response to both the 1998 NHIS interview and the MEPS interviews for rounds 1-3.

3.1.2 Panel 5

For MEPS Panel 5 5,357 households were fielded in 2000, a nationally representative subsample of the households responding to the 1999 National Health Interview Survey (NHIS). As described above for all previous MEPS samples, the Panel 5 sample reflects the oversampling of Hispanics and Black households for the NHIS.

The overall MEPS Panel 5 response rate at the end of round 1 (which collects data for the first part of 2000) was 74.9 percent. This overall rate reflects response to the 1999 NHIS interview and the MEPS round 1 interview.

3.1.3 Combined Panel Response

A pooled response rate for the survey respondents in this data set can be obtained by taking an average of the panel specific response rates. This pooled response rate for the combined panels is 70.0 percent.

3.2 Sample Weights

The sample weights provided in this file can be used to produce estimates for the U.S. civilian, non-institutionalized population and subgroups of this population based on the sample data. Two weights are provided: a person level weight and a family level weight.

3.2.1 Person Level Weight

The person level weight variable (WGTSP13) was constructed as a composite of separate panel specific weights. A positive person level weight was assigned to all key members of the U.S. civilian non-institutionalized population for whom MEPS data were collected, representing the corresponding U.S. population in early 2000. For the Panel 5, Round 1 participants this weight reflects the original household probability of selection for the NHIS, ratio-adjustment to NHIS national population estimates at the household level, adjustment for non-participation in MEPS at the household level, and poststratification to figures obtained from March 2000 Current Population Survey (CPS) data at the family and person levels. For Panel 4, Round 3 participants the weights also incorporate adjustments for nonresponse over the course of the year 1999. The person level poststratification reflected population distributions across census region, MSA status, race/ethnicity (Hispanic, black/non-Hispanic, other), sex, and age.

Overall, the weighted population estimate based on WGTSP13 for the civilian noninstitutionalized population is 273,309,558. Estimates can be made for this population based on the 25,094 sample persons in the file with positive weights (WGTSP13>0).

3.2.2 Family Level Weight

3.2.2.1 Definition of MEPS Families

A family unit is defined in MEPS as two or more persons living together in the same household during the reference period (in this data set, from January 1, 2000 to the date of interview) who are related by blood, marriage, or adoption (including foster children). In addition, unrelated persons who identify themselves as a family (e.g. domestic partners) are also defined as a MEPS family unit. Persons who died during the round 1 reference period and those who left the civilian non-institutionalized population part way through the reference period due to institutionalization, emigration, or enrollment in the military were considered to be family members. Relatives identified as usual residents of the household who were not there at the time of the interview, such as college students living away from their parents' home during the school year, were considered as members of the family that identified them.

3.2.2.2 Assignment of Weights

If all key in-scope members of a family responded to MEPS for their entire period of eligibility in 2000 for Round 3/Panel 4, or for Round 1/Panel 5, and the family had a key reference person, then that family received a family level weight (WGTRU13>0). Reporting units consisting of an individual respondent who was both key and in-scope also received a family level weight. These individual person units can be included or excluded from family level analyses at the analyst's discretion.

Family level weights were poststratified to figures obtained from the March 2000 CPS. The family level poststratification reflected population distributions across family type (reference person married, spouse present; male reference person, no spouse present; female reference person, no spouse present), size of family, age of reference person, location of family (census region and MSA status), and race/ethnicity of the family's reference person. The weighted estimate of the number of units (families plus individual person units) with family level weights containing at least one member of the U.S. civilian non-institutionalized population is 116,564,731, based on 9,927 families with WGTRU13>0.

It should be noted that CPS and MEPS definitions of family units are slightly different. In particular, CPS does not include foster children in families or consider unmarried persons who live together as family units. Adjustments were made in the poststratification process to help compensate for some of these differences.

3.2.2.3 Instructions to Create Family Estimates

To make estimates at the family level, it is necessary to prepare a family level file containing one record per family. Each MEPS family unit is uniquely identified by the combination of the variables DUID and FAMID13. Only persons with positive nonzero family weight values (WGTRU13>0) are candidates for inclusion in family estimates. Following is a summary of steps that can be used for family level estimation.

1. Concatenate the variables DUID and FAMID13 into a new variable (e.g. DUFAM13).

2. To create a family level file, sort by DUFAM13 and then subset to one record per DUFAM13 value by retaining only the reference person record (RNDREF13=1) for each value of DUFAM13. If the analyst chooses to eliminate single person units from family analyses, it is also necessary to exclude records where FAMSIZ13=1. If aggregate measures for families are needed for analytic purposes (e.g. means or totals), then those measures need to be computed using person-level information within families and attached to the family record. For other types of variables, analysts frequently use characteristics of the reference person to represent family characteristics.

3. Apply the weight WGTRU13 to the analytic variable(s) of interest to obtain national family estimates.

3.2.3 Relationship Between Person and Family Level Weights

Some persons with positive person level weights do not have family level weights because at least one member of their family was a non-participant in MEPS. Others with positive person weights did not receive a family weight because the family reference person was not key. In addition, some persons with positive family level weights do not have person level weights because they were either non-key or a member of the military during the first half of 2000. Analysts should include only persons with positive person level weights for person level analyses and persons with positive family level weights for family level analyses.

3.3 Variance Estimation

To obtain estimates of variability (such as the standard error of sample estimates or corresponding confidence intervals) for estimates based on MEPS survey data, one needs to take into account the complex sample design of MEPS for both person and family level analyses. Various approaches can be used to develop such estimates of variance including a Taylor series method for variance estimation or various replication methodologies. Replicate weights have not been developed for the MEPS data. We will describe the variables needed to implement a Taylor series estimation approach.

Using a Taylor Series approach, variance estimation strata and the variance estimation PSUs within these strata must be specified. The variables VARST13 and PSU13 on this MEPS data file (updated versions of corresponding variables provided in previously released MEPS public use files) serve to identify the sampling strata and primary sampling units required by the variance estimation programs. Specifying a "with replacement" design in a computer software package such as SUDAAN should provide estimated standard errors appropriate for assessing the variability of MEPS survey estimates. It should be noted that the number of degrees of freedom associated with estimates of variability indicated by such a package may not appropriately reflect the number available. For variables of interest distributed throughout the country (and thus the MEPS sample PSUs), one can generally expect to have at least 60 degrees of freedom associated with the estimated standard errors for national estimates based on this MEPS database.

D. VARIABLE-SOURCE CROSSWALK

SURVEY ADMINISTRATION VARIABLES

VARIABLE	LABEL	SOURCE
DUID	DU ID	Assigned in Sampling
PID	Person Number	Assigned in Sampling or by CAPI
DUPERSID	Person ID (DUID+PID)	Assigned in Sampling
PANEL13	Panel Number	Assigned by CAPI
FAMID13	Family Identifier (Student Merged In)	CAPI Derived
RULETR13	RU Letter	CAPI Derived
RUSIZE13	RU Size	CAPI Derived
RUCLAS13	RU Fielded As: Standard, New, Student	CAPI Derived
FAMSIZ13	RU Size Including Students	CAPI Derived
REGION13	Census Region	Assigned in Sampling
MSA13	MSA	Assigned in Sampling
RNDREF13	Reference Person	RE 42-45
RDRESP13	1st Respondent Indicator	RE 6, 8
PROXY13	Was Respondent A Proxy	RE 2
BEGRFD13	Reference Period Begin Date: Day	CAPI Derived
BEGRFM13	Reference Period Begin Date: Month	CAPI Derived
BEGRFY13	Reference Period Begin Date: Year	CAPI Derived
ENDRFD13	Reference Period End Date: Day	CAPI Derived
ENDRFM13	Reference Period End Date: Month	CAPI Derived
ENDRFY13	Reference Period End Date: Year	CAPI Derived
KEYNESS	Person Key Status	RE Section
INSCOP13	In-scope	RE Section
PSTAT13	Person Disposition Status	RE Section
RURSLT13	RU Result	Assigned by CAPI
RUENDD13	Date of Intv (Date Started: Day)	Assigned by CAPI
RUENDM13	Date of Intv (Date Started: Month)	Assigned by CAPI
RUENDY13	Date of Intv (Date Started: Year)	Assigned by CAPI

DEMOGRAPHIC VARIABLES

VARIABLE	LABEL	SOURCE
AGE13X	Age - (Edited/Imputed)	RE 12, 57-66
DOBMM	Date of Birth: Month	RE 12, 57-66
DOBYY	Date of Birth: Year	RE 12, 57-66
SEX	Sex	RE 12, 57, 61
RACEX	Race - (Edited/Imputed)	RE 101, 102
RACETHNX	Race/Ethnicity - (Edited/Imputed)	RE 98-102
HISPANX	Hispanic Ethnicity - (Edited/Imputed)	RE 98-100
HISPCAT	Specific Hispanic Ethnicity Group	RE 98-100
MARRY13X	Marital Status - (Edited/Imputed)	RE 13, 97
SPOUID13	Spouse ID	RE 13, 97
SPOUIN13	Marital Status W/ Spouse Present	RE 13, 97
EDUCYR13	Years of Educ when First Entered MEPS	RE 103-105
HIDEG13	Highest Degree when First Entered MEPS	RE 103-105
FTSTD13X	Student Status Ages 17-23 (Edit/Imputed)	RE 11A, 106-108
ACTDTY13	Military Full-Time Active Duty	RE14, 96
RFREL13X	Relation To Ref Pers (Edited/Imputed)	RE 76-77

HEALTH STATUS VARIABLES

VARIABLE	LABEL	SOURCE
RTHLTH13	Perceived Health Status	CE 1
RTPROX13	Self/Proxy Rating of Health	CE 1
MNHLTH13	Perceived Mental Health Status	CE 2
MNPROX13	Self/Proxy Rating of Mental Health	CE 2
IADLHP13	IADL Screener	HE 2,3
ADLHLP13	ADL Screener	HE 5,6
AIDHLP13	Uses Assistive Devices	HE 7,8
WLKLIM13	Limitation in Physical Functioning	HE 9,10
LFTDIF13	Difficulty Lifting 10 Pounds	HE 11
STPDIF13	Difficulty Walking Up 10 Steps	HE 12
WLKDIF13	Difficulty Walking 3 Blocks	HE 13
MILDIF13	Difficulty Walking a Mile	HE 14
STNDIF13	Difficulty Standing 20 Minutes	HE 15
BENDIF13	Difficulty Bending/Stooping	HE 16
RCHDIF13	Difficulty Reaching Over Head	HE 17
FNGRDF13	Difficulty Using Fingers to Grasp	HE 18
ACTLIM13	Limitation Work/Housework/School	HE 19,20
WRKLIM13	Work Limitation	HE 19,20
HSELIM13	Housework Limitation	HE 19,20
SCHLIM13	School Limitation	HE 19,20
UNABLE13	Completely Unable To Do Activity	HE 21
SOCLIM13	Social Limitation	HE 22,23
COGLIM13	Cognitive Limitation	HE 24,25

EMPLOYMENT VARIABLES

VARIABLE	LABEL	SOURCE
EMPST13	Employment Status	EM 1-3; RJ 1, 6
HRWAG13X	Hourly Wage at Current Main Job	EW section EM 104, 111
HRWGRD13	Hourly Wage Round Flag	Constructed.
HRWAY13	Calculation Methods for Hourly Wage	EM 1-3, 51, 65,104, 111; EW section
HOUR13	Hours Worked Per Week at CMJ	EM 1-3, 51, 65, 104-105, 111; EW 17; RJ 1
HELD13X	Health Insurance Held From CMJ	EM, HX, RJ and HP sections
OFFER13X	Health Insurance Offered at CMJ	EM, HX, RJ and HP sections
NUMEMP13	Number of Employees at Location of CMJ	EM 91-92, 124; RJ 8
SELFCM13	Self-Employed at Current Main Job	EM 1-3, 5, 11, 18, 27, 40, 53; RJ 1, 6

HEALTH INSURANCE VARIABLES

VARIABLE	LABEL	SOURCE
CHNOW13X	PID Cov By CHAMPUS/CHAMPVA/TRICARE at Int Date - Edited	HX 12, 13; PR 19 - 22; HQ section; AGE13X; RE 14, 96A
MCARE13	PID Cov By MEDICARE	HX 5 - 7
MCARE13X	PID Cov By MEDICARE - Edited	HX 5 - 7, 10 - 15; PRIV13 and (HX 48 or (OE 10, 24, 37)); PR 7-10, 19-26
OTPUB13X	PID Cov By Other Public Ins - Edited	HX 10, 11, 14, 15, 18, 19; HQ section; PR 7-10, 23-26, 39-42
PRIV13	PID Cov By Private Ins	HX 2 - 4, 21 - 24, 48; HP, OE, HQ, EM and RJ sections
INSRD13X	PID Is Insured - Edited	CHNOW13X, MCARE13X, OTPUB13X, PRIV13

WEIGHTS

VARIABLE	LABEL	SOURCE
WGTSP13	Person Weight	Constructed
WGTRU13	Family Weight	Constructed
VARST13	Variance Estimation Stratum	Constructed
PSU13	Variance Estimation PSU	Constructed